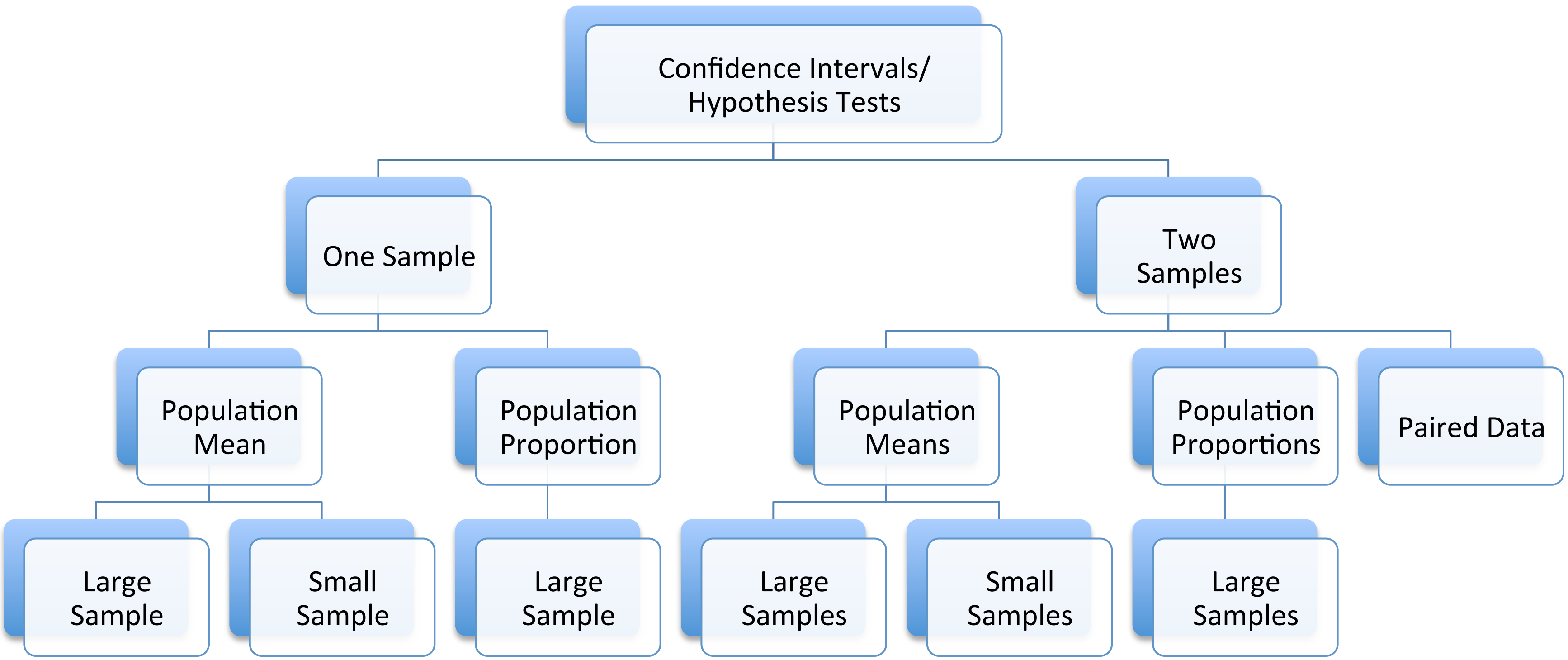
More Hypothesis Testing: Difference in Proportions and Paired Data

Keegan Korthauer
Department of Statistics
UW Madison

Outline



Confidence Intervals/
Hypothesis Tests
One Sample
Two Samples
Population Mean
Population Proportion
Population Means
Population Proportions
Paired Data
Large Sample
Small Sample
Large Sample
Large Samples
Small Samples
Large Samples

Motivation

- Recall the example: A randomized double-blind experiment assigned 244 smokers to receive nicotine patches and another 245 to receive patches and an antidepressant. After a year, 40 in the first group and 87 in the second had quit.

- Can we conclude that the proportion of smokers who quit is different in the two groups?

The Idea

- One-sample test statistic (population proportion):
  - observed $\hat{p}$ too far from p_0 -> reject H_0
  - observed $\hat{p}$ close to p_0 -> do not reject H_0

- Two-sample test (difference in proportions):
  - observed difference $\hat{p}_X$ - $\hat{p}_Y$ too far from 0 -> reject H_0 (proportions are not equal)
  - observed difference $\hat{p}_X$ - $\hat{p}_Y$ close to 0 -> do not reject H_0 (proportions are equal)

→ Construct a test statistic using $\hat{p}_X$ - $\hat{p}_Y$

Deriving the Test Statistic

- Recall with large samples, the CLT (normal approximation to the binomial) gives us

$$\hat{p}_X \sim N\left(p_X,\ \frac{p_X(1-p_X)}{n_X}\right) \text{ and } \hat{p}_Y \sim N\left(p_Y,\ \frac{p_Y(1-p_Y)}{n_Y}\right)$$

- Combining these, we get the following results, which will come in handy when we compute the p-value

$$\hat{p}_X - \hat{p}_Y \sim N\left(p_X - p_Y,\ \frac{p_X(1-p_X)}{n_X} + \frac{p_Y(1-p_Y)}{n_Y}\right), \text{ so}$$

$$\frac{(\hat{p}_X - \hat{p}_Y) - (p_X - p_Y)}{\sqrt{p_X(1-p_X)/n_X + p_Y(1-p_Y)/n_Y}} \sim N(0,1)$$

Deriving the Test Statistic - continued

- Under the null hypothesis that there is no difference:

$$p_X = p_Y \Rightarrow p_X - p_Y = 0$$

- Best guess for the common population proportion is the pooled proportion:

$$\hat{p} = \frac{X + Y}{n_X + n_Y}$$

$$\text{Estimate } \frac{p_X(1 - p_X)}{n_X} + \frac{p_Y(1 - p_Y)}{n_Y} \text{ with } \hat{p}(1 - \hat{p})\left(\frac{1}{n_X} + \frac{1}{n_Y}\right)$$

- Which gives us a z-score:

$$z = \frac{\hat{p}_X - \hat{p}_Y}{\sqrt{\hat{p}(1 - \hat{p})(1/n_X + 1/n_Y)}}$$

HT for Difference in Proportions

Let X ~ Bin(n_X, p_X) and Y ~ Bin(n_Y, p_Y). Assume there are at least 10 successes and 10 failures in each sample and that X and Y are independent

1. Set up the null H_0 and alternative H_1 hypotheses (see table below)
2. State the level of significance α you will use
3. Calculate the **z-score** (test statistic):
   where $\hat{p}$ is the pooled proportion

$$z = \frac{(\hat{p}_X - \hat{p}_Y)}{\sqrt{\hat{p}(1-\hat{p})(1/n_X + 1/n_Y)}}$$

4. Assume H_0 is true and calculate the P-value:

H_0	H_1	P-value
$p_X - p_Y \leq 0$	$p_X - p_Y > 0$	Area to the right of *z*
$p_X - p_Y \geq 0$	$p_X - p_Y < 0$	Area to the left of *z*
$p_X - p_Y = 0$	$p_X - p_Y \neq 0$	Area to the left of -z plus area to the right of z

5. Make a conclusion based on the P-value

Note on Sample Size

- Since the normal approximation to the binomial relies on the CLT, we have to have some restriction on the sample size

- The HT given on the previous slide is valid when there are at least 10 successes and failures in each sample:

$$X \geq 10 \text{ and } n_X - X \geq 10 \text{ and}$$

$$Y \geq 10 \text{ and } n_Y - Y \geq 10$$

Example – Smoking Cessation

- Recall the example: A randomized double-blind experiment assigned 244 smokers to receive nicotine patches and another 245 to receive patches and an antidepressant. After a year, 40 in the first group and 87 in the second had quit.

- Can we conclude that the proportion of smokers who quit is different in the two groups?

TWO-SAMPLE TESTS FOR PAIRED DATA

Motivation

Recall the gas mileage example:

- Each of 10 cars was first filled with either regular or premium gas, decided by a coin toss, and the mileage for that tank was recorded.
- The mileage was recorded again **for the same cars** using the other kind of gasoline.

Car	1	2	3	4	5	6	7	8	9	10
Premium	19	22	24	24	25	25	26	26	28	32
Regular	16	20	21	22	23	22	27	25	27	28
Difference	3	2	3	2	2	3	-1	1	1	4

- Can we conclude that the mean gas mileage for these cars is greater when using premium gasoline than when using regular?

Deriving the Test Statistic for Paired Data

- Let $(X_1,Y_1),\ldots,(X_n,Y_n)$ be the n **paired** observations
  - for example, X_i is the gas mileage on premium for the i^{th} car and Y_i is the gas mileage on regular for the i^{th} car
- Then let $D_i = X_i - Y_i$
  - for example, D_i is the difference in gas mileage for premium and regular gasoline for the i^{th} car
- We can treat the single sample of differences D_i as single random sample from a population of differences with mean μ_D and standard deviation σ_D
- Then we can use the Hypothesis Test for a Population Mean (for large (6.2) or small samples (6.4) depending on the value of n)

Test for Paired Data (Small Sample)

Let $D_1,\ldots,D_n$ be a **small** random sample of $n \leq 30$ differences of pairs that follow a **normal** distribution with mean μ_D (unknown sd σ_D).

1. Set up the null H_0 and alternative H_1 hypotheses (see table below)
2. State the level of significance α you will use
3. Calculate the test statistic: $$t = \frac{\bar{D} - \mu_0}{s_D / \sqrt{n}}$$
4. Assume H_0 is true and calculate the P-value using areas under the t curve with n-1 degrees of freedom:

H_0	H_1	P-value
$\mu_D \leq \mu_0$	$\mu_D > \mu_0$	Area to the right of t
$\mu_D \geq \mu_0$	$\mu_D < \mu_0$	Area to the left of t
$\mu_D = \mu_0$	$\mu_D \neq \mu_0$	Area to the left of -t plus area to the right of t

5. Make a conclusion based on the P-value

Test for Paired Data (Large Sample)

- Same procedure as with a small sample, except
  - Do not need to assume normal population of differences
  - The test statistic will be a z-score instead of a t-statistic
  - P-values calculated with Standard Normal Table/pnorm() R function instead of t Table/pt() R function

Example

Recall the gas mileage example:

- Each of 10 cars was first filled with either regular or premium gas, decided by a coin toss, and the mileage for that tank was recorded.

- The mileage was recorded again **for the same cars** using the other kind of gasoline.

Car	1	2	3	4	5	6	7	8	9	10
Premium	19	22	24	24	25	25	26	26	28	32
Regular	16	20	21	22	23	22	27	25	27	28
Difference	3	2	3	2	2	3	-1	1	1	4

- Can we conclude that the mean gas mileage for these cars is greater when using premium gasoline than when using regular?

Next

- Nonparametric (Distribution-Free) tests
- Chi-Square Test
- F Test
- Power & Type I Error